Approaches to improve automation for security

Sara Matzner
Program Manager,
Cyber Information Assurance & Decision Support (CIADS)
Information Systems Laboratory
Applied Research Laboratories
The University of Texas at Austin
matzner@arlut.utexas.edu, 512-835-3176

CIADS domain of expertise

- Information assurance
  - Telecommunications and computer networks
- Expert systems for intrusion detection
- Vulnerability assessment
- Network modeling and simulation

Problem Statement

- Networks are vulnerable.
  - External and internal sources of threat
- Intrusion detection systems are imperfect.
  - High false alarm rates
- Threat assessment is manpower-intensive.
  - Overwhelming quantity of data

Goals

- **Support the analyst using state of the art technologies**
- **Provide decision support through data management**
  - Data reduction,correlation,summarization
- **Provide both post-analysis and real time response capabilities**
- **Bridge policy and compliance**
  - Dynamic policy updates
- **Automate detection tasks where possible**

Strategy for near-term

Funding needed:

- Extension of current technological approaches
- Techniques for automation are coming to maturity now

Techniques for automation

- Machine learning
  - Developed through data mining of historical databases
- Artificial intelligence
  - Autonomous agents, genetic algorithms, neural networks
- Payoff: automation and extension of human pattern recognition capabilities

Data Mining

- Knowledge discovery in databases using:
  - Clustering
  - Classification
  - Association Rule Mining
  - High-Dimensional Visualization
- Benefits:
  - Discovery of attack sequences
  - Characterization of normal conditions in order to recognize abnormal behavior
  - Represents current state-of-the-art

Artificial Intelligence

- Autonomous Agents
  - Actively gather data as needed
    - Confirmatory Agents: Used to fill in gaps in data-mining-based hypotheses concerning intrusions
    - Discovery Agents: Used to find anomalous situations

Artificial Intelligence

- Autonomous Agents
  - Example uses:
    - Vulnerability analysis: “automated Red Team”
      - Coupled with genetic algorithms to randomize attack sequences
    - Data retrieval: an agent to penetrate hostile and friendly systems
    - Countermeasure deployment: a means to compromise a target system

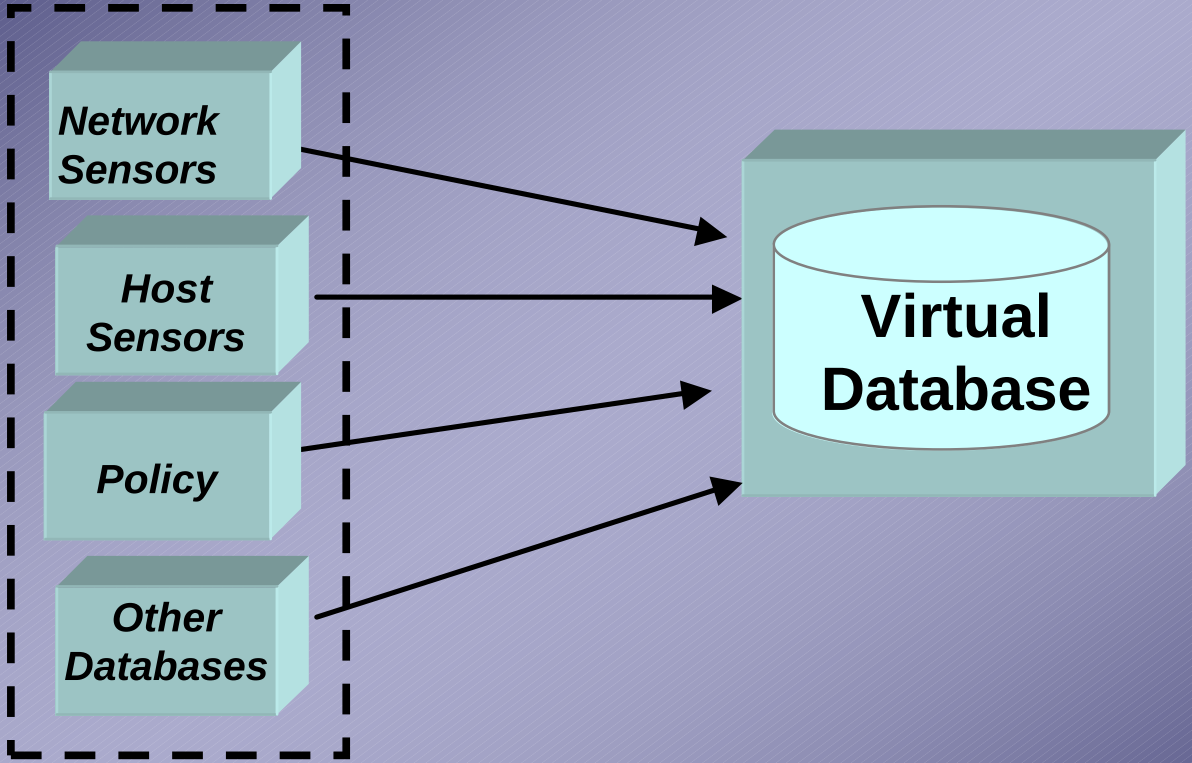
Status

- Knowledge Engineering & Data Mining
  - Capture what you know (but don't know you know)
  - Discovery of new relations in existing data
  - Represents current technology
  - Currently performed offline (post analysis)
  - Remain fairly human intensive

Automated Data Retrieval

CIADS

Applied Research Laboratories, The University of Texas at Austin



Changing environment

- Computing environment is becoming more distributed and changing dynamically
  - Data, processing and knowledge will be distributed throughout the network
    - Distributed knowledge will allow for recognizing correlations across broad regions of the network.
    - Data analysis and filtering will occur at lower-levels
      - Caveat – Information will not be available for higher-level synthesis
  - Network topology will change in a shortened time scale

Impact

- Greater analysis load on the human
- Requires more synthesis of information and more automation at all levels